

Projekt 1 (Python)

Maksymalna liczba punktów do uzyskania: 10. Można zrobić jedno lub dwa zadania.

Projekt polega na przygotowaniu pliku o nazwie `Imię_Nazwisko_album_Python_1.ipynb` (np. `Wiktor_Wojciechowski_675467_Python_1.ipynb`) z kodem napisanym w języku Python. Oprócz funkcji, której stworzenie jest przedmiotem każdego zadania, należy także przedstawić jeden przykład ilustrujący jej zastosowanie.

Do wykonania projektu potrzebne są dane z następujących plików:

- `wybory_sejm_2015.csv`, `wybory_sejm_2019.csv`, `wybory_sejm_2023.csv`
- `bezrobocie.xlsx` (proszę w kodzie uwzględnić pobieranie danych odpowiedniego arkusza)
- `wynagrodzenia.xlsx` (proszę w kodzie uwzględnić pobieranie danych odpowiedniego arkusza)

Plik z wynikami wyborów w 2023 roku zawiera osobno dane o wyborach przeprowadzonych zagranicą. Dla uproszczenia zadania, na potrzeby niniejszego projektu, proszę te głosy zignorować.

Zadanie 1 [6 pkt]

Napisz funkcję `election_statistics`, która będzie tworzyła obiekt typu Data Frame z jedną z następujących statystyk:

- Frekwencja ogółem (liczba kart wyjętych z urn/ liczba osób uprawnionych do głosowania) ('frekwencja')
- Udział głosów nieważnych (chodzi o nieważne głosy, a nie karty) w całkowitej liczbie oddanych głosów ('głosy nieważne')
- Udział wyborców głosujących na podstawie zaświadczenia o prawie do głosowania w całkowitej liczbie oddanych głosów ('zaświadczenia')

Funkcja `election_statistics` powinna przyjmować następujące argumenty:

- `year` = lista, domyślnie `[2015, 2019, 2023]`
- `jst` – string określający czy funkcja ma przedstawiać statystyki na poziomie gminy, powiatu czy województwa (tj, 'gmina', 'powiat', 'województwo'), domyślnie funkcja powinna generować statystyki na poziomie powiatu,
- `jst_number` – domyślna wartość 10,
- `sort` – domyślna wartość logiczna `True` dla sortowania od wartości najmniejszej do największej (`False` dla sortowania od wartości największej do najmniejszej),
- `sort_year` = numer indeksu z listy `year`, domyślnie ostatnia element listy,
- `stat` = string określający wybraną statystykę: 'frekwencja', 'głosy nieważne', 'zaświadczenia'

Uwagi do funkcji:

- Nazwy gmin i powiatów nie powinny zawierać słów/skrótów typu 'gm.', 'powiat'. Czyli np. zamiast 'gm. Węgorzyno' powinno być 'Węgorzyno',
- Formatowanie wartości procentowych: 20, a nie 0,2 lub 20%,
- Funkcja powinna zawierać docstring z wyczerpującymi informacjami do czego służy, jakie może przyjmować argumenty, itp.,
- W przypadku wskazania liczby regionów do porównania, które wykraczają poza liczbę istniejących regionów, funkcja powinna zwracać komunikat o błędzie.

Przykład zastosowania:

```
election_statistics(year = [2015, 2019], jst='powiat', jst_number=8, sort=True, stat='frekwencja')
```

Powyższa funkcja powinna zwrócić ramkę danych z nazwami 8 powiatów (dodatkowo także w osobnej kolumnie nazwę województwa), w których w 2019 roku odnotowano najniższą frekwencję w wyborach do Sejmu łącznie z poziomem tej frekwencji w 2015 i 2019 roku.

```
election_statistics(year = [2015, 2023], jst='gmina', jst_number=8, sort=False, stat='zaświadczenia')
```

Powyższa funkcja powinna zwrócić ramkę danych z nazwami 8 gmin (dodatkowo także w osobnych kolumnach nazwy powiatu i województwa), w których w 2023 roku odnotowano najwyższy udział osób głosujących na podstawie zaświadczeń o prawie do głosowania w wyborach do Sejmu łącznie z wysokością tych udziałów w 2015 i 2023 roku.

Zadanie 2 [4 pkt]

Napisz funkcję *election*, która będzie tworzyła obiekt Data Frame porównujący przeciętne wartości poniższych statystyk we wskazanych latach wyborczych dla *n* powiatów uszeregowanych według łącznej liczby oddanych głosów (malejąco):

- procentowe poparcie dla poszczególnych obozów politycznych (poparcie liczymy jako procentowy udział oddanych głosów na dany obóz polityczny),
- stopa bezrobocia rejestrowanego,
- przeciętne wynagrodzenie brutto (przyjmij, że wynagrodzenie w 2023 roku jest takie samo, jakie było w 2022 roku).

Funkcja *election_data* powinna przyjmować następujące argumenty:

- year = lista, domyślnie [2015, 2019, 2023]
- parties = lista stringów z nazwami obozów politycznych, domyślnie ['PiS', 'KO', 'PSL', 'Lewica', 'Konfederacja', 'Mniejszość niem.'],
- count – liczba określająca liczbę powiatów uwzględnionych w analizie, domyślnie 1,

Uwagi do funkcji:

- Formatowanie wartości procentowych: 5, a nie 0,05 lub 5%,
- Nazwy powiatów nie mogą zawierać słów/skrótów typu 'm.', 'powiat'. Czyli np. zamiast 'm. Szczecin' powinno być 'Szczecin',
- Funkcja powinna zawierać docstring z wyczerpującymi informacjami do czego służy, jakie może przyjmować argumenty, itp.,
- Partie polityczne/komitety, które należy wliczyć do poszczególnych obozów politycznych zostały przedstawione poniżej.

Przykład zastosowania:

```
election (year = [2015, 2019], parties = ['PiS', 'KO', 'PSL'], count = 2)
```

Powyższa funkcja powinna zwrócić ramkę danych pokazującą przeciętne statystyki w dwóch powiatach o największej łącznej liczbie oddanych głosów oddzielnie w każdym województwie i roku. Ramka powinna być posortowana według nazwy województw i lat (nie odwrotnie). Statystyki powinny dotyczyć procentowego poparcie dla 'PiS', 'KO', 'PSL' (liczonego jako suma głosów oddanych na każdy z

obozów podzielona przez sumę łącznej liczby ważnych głosów, a nie jako średnia arytmetyczna) oraz przeciętnej stopy bezrobocia (średnia arytmetyczna) i wynagrodzenia w tych powiatach (średnia arytmetyczna).

Na potrzeby projektu do poszczególnych obozów politycznych należy wliczyć oddane głosy według poniższej instrukcji (nazwy kolumn w odpowiednich plikach):

PiS

- KOMITET WYBORCZY PRAWO I SPRAWIEDLIWOŚĆ (2023),
- KOMITET WYBORCZY PRAWO I SPRAWIEDLIWOŚĆ - ZPOW-601-9/19 (2019),
- 1 - Komitet Wyborczy Prawo i Sprawiedliwość (2015)

KO

- KOALICYJNY KOMITET WYBORCZY KOALICJA OBYWATELSKA PO .N IPL ZIELONI (2023),
- KOALICYJNY KOMITET WYBORCZY KOALICJA OBYWATELSKA PO .N IPL ZIELONI - ZPOW-601-6/19 (2019),
- 2 - Komitet Wyborczy Platforma Obywatelska RP (2015) oraz 8 - Komitet Wyborczy Nowoczesna Ryszarda Petru (2015)

PSL

- KOALICYJNY KOMITET WYBORCZY TRZECIA DROGA POLSKA 2050 SZYMONA HOŁOWNI - POLSKIE STRONNICTWO LUDOWE (2023),
- KOMITET WYBORCZY POLSKIE STRONNICTWO LUDOWE - ZPOW-601-19/19 (2019)
- 5 - Komitet Wyborczy Polskie Stronnictwo Ludowe (2015)

Konfederacja

- KOMITET WYBORCZY KONFEDERACJA WOLNOŚĆ I NIEPODLEGŁOŚĆ (2023),
- KOMITET WYBORCZY KONFEDERACJA WOLNOŚĆ I NIEPODLEGŁOŚĆ - ZPOW-601-5/19 (2019),
- 4 - Komitet Wyborczy KORWiN (2015) oraz 14 - Komitet Wyborczy Wyborców Grzegorza Brauna „Szczęść Boże!” (2015)

Lewica

- Lewica - KOMITET WYBORCZY NOWA LEWICA (2023),
- KOMITET WYBORCZY SOJUSZ LEWICY DEMOKRATYCZNEJ - ZPOW-601-1/19 (2019)
- 3 - Komitet Wyborczy Partia Razem (2015) oraz 6 - Koalicyjny Komitet Wyborczy Zjednoczona Lewica SLD+TR+PPS+UP+Zieloni (2015)

Mniejszość niemiecka

- KOMITET WYBORCZY WYBORCÓW MNIEJSZOŚĆ NIEMIECKA (2023),
- KOMITET WYBORCZY WYBORCÓW MNIEJSZOŚĆ NIEMIECKA - ZPOW-601-15/19 (2019),
- 16 - Komitet Wyborczy Wyborców Mniejszość Niemiecka (2015)